I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 16225

Are Informal Self-Employment and Informal Employment as Employee Behaviorally Distinct Labor Force States?

Luca Flabbi
Mauricio Tejada

JUNE 2023

DISCUSSION PAPER SERIES

IZA DP No. 16225

Are Informal Self-Employment and Informal Employment as Employee Behaviorally Distinct Labor Force States?

Luca Flabbi
University of North Carolina and IZA

Mauricio Tejada
Universidad Diego Portales

JUNE 2023

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Are Informal Self-Employment and Informal Employment as Employee Behaviorally Distinct Labor Force States?*

The paper performs both a parametric and non-parametric analysis to address a fundamental question in the growing literature using search models to study labor market informality: should informal self-employment and informal employment as employee be considered two different labor market states? Both the non-parametric and the parametric tests strongly reject equality between the two states, cautioning against aggregating them in a common "informality state." The parametric model indicates the source of the difference in the high dispersion of informal self-employment income and in the low duration of informal employee jobs.

JEL Classification: J46, J64, O17

Keywords: labor market frictions, search and matching, informality, self-employment

Corresponding author:
Luca Flabbi
Economics Department
University of North Carolina at Chapel Hill
107 Gardner Hall, CB 3305
Chapel Hill, NC 27599-3305
USA
E-mail: luca.flabbi@unc.edu

* Tejada gratefully acknowledges financial support from FONDECYT, grant project No. 11196296

1 Introduction

In an influential paper published in 1983, Flinn and Heckman asked: "Are Unemployment and out of the Labor Force Behaviorally Distinct Labor Force States?" (Flinn and Heckman, 1983). The question was relevant because labor economists had started to study labor market dynamics with richer theoretical and empirical models, forcing researchers to take a stand on which labor market states were relevant. The study of labor market dynamics in economies with high informality is experiencing a similar transformation. Richer labor market models have recently been developed and estimated, prompting a crucial debate on what the relevant labor market states are, and which transitions between them we should focus on.

This paper contributes to the debate by addressing its most controversial question: should we differentiate between informal workers that are hired as employee and those that are working as self-employed? Aggregating or differentiating these two labor market states has proven relevant both for obtaining credible estimates and for drawing policy implications but both approaches have been used by influential papers, without producing a consensus in the literature. For example, Meghir et al. (2015) is one of the first estimated search model on a market with high informality (Brazil) and aggregates all the unregistered employees and the self-employed in the same labor market state; Bobba et al. (2022) is a recent contribution on Mexico but strongly differentiates between the two states, so much so as to consider self-employed informality as a searching state in alternative to unemployment. A number of other contributions take one approach or the other.[1]

We use data from Colombia – the fourth economy in Latin America, a region with particularly high levels of labor market informality – to test if the informal self-employed and the informal employee should belongs to two different labor market states when modeling labor market dynamics. We follow Flinn and Heckman (1983) in providing two types of analysis. First, we conduct

[1]Among the contributions that do *not* differentiate between informal employee and self-employed are early contributions in the theoretical search literature, such as Albrecht et al. (2009); Charlot et al. (2013); and in the macro search literature, such as Bosch and Esteban-Pretel (2012). More recently, Haanwinckel and Soares (2021) develop a search model with intra-firm bargaining and exclude the self-employed. On the other side, contributions developing search models of the labor market that *do* differentiate informal self-employment as a distinct state from informal employment include Bobba et al. (2021); Narita (2020).

Other examples beyond the search literature that take a stand in this debate include: Esteban-Pretel and Kitao (2021), which allows for only one informal labor market state, excluding the self-employed from the calibration sample; Ulyssea (2018), which considers informality choices of both firms and workers, separating informal workers who are employees from informal firms; Granda and Hamann (2015), which distinguishes between informal entrepreneur (informal self-employed) and informal worker (informal employee); Almeida and Carneiro (2012), which clearly differentiates between informal wage earners and self-employed.

non-parametric tests for equality of the empirical duration and labor income distributions. Then, we develop and estimate a search model for Colombia where we can directly impose the same behavior for informal self-employed and informal employee and test the restrictions with likelihood ratio tests.

Both the non-parametric and the parametric tests strongly reject that informal self-employment and informal employment as employees are behaviorally indistinguishable labor force states. The result cautions against aggregating them when studying labor markets with high informality or using only one of the two as representative of the the typical informal worker. We estimate the main sources of the difference to be the dispersion in labor income offers (much higher for the informal self-employed) and the job termination rate (much higher for the informal employees).

The paper is organized as follows: Section 2 presents the data, Section 3 provides the non-parametric analysis and Section 4 the parametric one. Section 5 concludes.

2 Data

We use the Colombian *Gran Encuesta Integrada de Hogares (GEIH)* for 2016. GEIH is a nationally representative survey collected monthly by the *Administrative Department of National Statistics* (DANE). The survey contains individual characteristics – such as gender, age, and schooling – and collects labor market outcomes – such as employment status, durations, monthly labor income, weekly hours worked, and occupational characteristics. It does also allow for a precise definition of labor market informality. We define any employed workers to be *informal* if they do not contribute to social security, a definition consistent with International Labor Organization (ILO)'s recommendations and with the previous literature on LAC (Perry et al., 2007; Kanbur, 2009; Bobba et al., 2022). If these workers are in a subordinate working relationship with a well-defined employer, we classify them as *informal employee*; if they are occupied in an activity with more independence and autonomy so that they declare themselves self-employed,[2] we classify them as *informal self-employed*. The typical activity is a small informal selling point in a city street corner.

To be consistent with the theoretical model, we extract an estimation sample relatively homogeneous over demographic characteristics and education. We focus on 25-55 years old men, living in urban areas, that have completed at most secondary education and work full-time when employed. We focus on male unskilled workers because this is the group on which labor market

[2] The original Spanish in the questionnaire is *trabajador por cuenta propia*, which means self-employed who is not an employer, i.e. that works on his own without employing other workers as subordinates.

informality is the most relevant and the most studied. To gain sample size, we pool together all the surveys conducted in the year, from January to December 2016. The final estimation sample include 88,123 observations, of which about 9% are unemployed, 51% informal – 39% self-employed and 12% employee –, and 40% are formal employees. Two important differences between informal employees and informal self-employed emerge from simple descriptive statistics of the estimation sample. First, the labor income distribution for the self-employed is much more dispersed than the one for informal employees, with standard deviations of, respectively, 0.56 and 0.36 (to be compared with means of 1.073$/h and 1.064$/h). Second, informal self-employed jobs last much longer than informal employee jobs, with an average of, respectively, 106 and 33 months. In the other two labor market states, the unemployed are searching for a job for an average of 4 months, while formal employees earn on average more than informal workers (with a mean of 1.419$/h). We should mention that it is possible to be a formal self-employed worker but no one is in this labor market state in this demographic group and therefore we ignored it in the analysis.

3 Non-parametric tests

To assess if informal self-employment and informal employment as employees are behaviorally indistinguishable states, we follow Flinn and Heckman (1983) and test if the distributions describing the two states are equal or not. If they are, they should not be considered separate states; if they are not, they should. Given the data at our disposal, we can non-parametrically estimate at least two distributions pertaining to each state. The first is the distribution of the duration a given worker stays in each of the two states. The second is the distribution of the wage (if informal employee) or self-employment income (if informal self-employed) a given worker receive in each of the two state. To these two distribution, we add a third distribution: the duration in the state of unemployment before transiting either to the informal employee state or to the informal self-employed state. While the first two distributions describe the state, the third capture factors that are leading the worker to one of the two states.

We non-parametrically estimate the duration distributions using the Kaplan-Meier survival function estimator[3] and estimate the labor income distribution using the empirical cumulative distribution function.[4] The estimated distributions are reported in Appendix Figures A.1 and A.2.

[3]That is $S(t_j) = \Pi_{i=1}^{j} \frac{n_i - h_i}{n_i}$ where t_j is the duration of the spell j, h_j is the number of completed spell of duration t_j, and n_j is the total number of spells not completed before t_j (Kiefer, 1988; Kaplan and Meier, 1958).

[4]The empirical cumulative distribution function is defined as $F(x_j) = \Pr[x_i \leq x_j] = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}[x_i \leq x_j]$ where $\mathbf{1}[C]$ is an indicator variables that takes the value of 1 es the condition C is satisfied and zero otherwise

Once estimated, we use the Kolmogorov-Smirnov (K-S) non-parametric to test for equality of the distributions. In the K-S test, the null hypothesis is whether the data draws composing the two observed samples come from the same underlying population distribution (Dodge, ed, 2008). Formally, the K-S test is based on the maximum difference between the empirical cumulative distribution functions of the two samples[5] and therefore does not provide information on how the two distributions differ, but only on whether they differ or not.

The test results are presented in Table 1. The null hypothesis of equal duration distributions between informal employees and informal self-employed is largely rejected as is the test of equal labor income distributions. The equality of unemployment duration distributions before transiting to either one or the other informal job state is also rejected but with a higher p-value, which corresponds to a 2% confidence level.

4 Parametric tests based on Search Model

The non-parametric tests presented in Section 3 already give a strong indication that informal employees and informal self-employed should be considered two separate labor market states. But they cannot distinguish if the sources of the separation are the frictions and shocks affecting the labor market dynamics or the wage offers distributions affecting agent's decisions to accept a job. To make progress in this understanding, we follow again Flinn and Heckman (1982) and develop and estimate a simple search model of the Colombian labor market. Under this parametric approach, we can directly impose the same behavior for informal self-employed and informal employee and perform likelihood ratio tests to assess the validity of the restriction.

4.1 Environment and equilibrium conditions

Time is continuous, the environment is stationary and the economy is populated by infinitely-lived individuals with discount rate ρ. Individuals can be in one of the following four states: unemployment u, informal self-employment s, formal employment f, and informal employment i. We denote with $v = s, f, i$ the different job type in which an agent can be employed in. While unemployed, individuals receive flow utility b and search for jobs, meeting offers at a Poisson rate $\lambda(v)$. Offers are fully described by labor income x, drawn from the exogenous offer distributions $G(x|v)$. All jobs terminate exogenously at rate $\eta(v)$. Formal employees are different from informal

and n is the total number of observations.

[5]The K-S test uses the empirical cumulative distribution function $F(x_j)$, as defined in footnote 4, for two samples, say 1 and 2, to compute the following statistic: $D = \sup_{x_j} |(F_1(x_j) - F_2(x_j))|$.

workers because they pay a proportional payroll contribution τ. In exchange, they receive benefits that are valued at a flow utility that we denote with θ. Only unemployed individuals search for a job.

We denote by U and $E(x,v)$ the steady-state values of unemployment and employment, respectively, leading to the following Bellman equation representation:

$$\rho U = b + \sum_{v=s,f,i} \lambda(v) \left[\int \max\{E(x,v) - U, 0\}\, dG(x|v) \right] \tag{1}$$

$$\rho E(x,v) = x\left[1 - \tau \iota_{v=f}\right] + \theta \iota_{v=f} + \eta(v)\left[U - E(x,v)\right] \tag{2}$$

where $\iota_{v=f}$ is an indicator variable equal 1 if the job is formal and zero otherwise.

The optimal decision for accepting a job offer possesses a reservation values property: the reservation labor income $x^*(v)$ for job v satisfies $E(x^*(v),v) = U$, leading to:

$$x^*(v) = \frac{\rho U - \theta \iota_{v=f}}{1 - \tau \iota_{v=f}} \tag{3}$$

which also implies $\theta = \tau \rho U$. In the steady-state equilibrium, flows are balanced, implying $u\lambda(v)\left[1 - G(x^*(v)|v)\right] = \eta(v)v$.

In this model environment, the *null hypothesis* of informal self-employment and informal employment as employee being behaviorally indistinguishable labor market states is represented by the following set of parametric constraints:

$$\begin{cases} \lambda(s) = \lambda(i) \\ \eta(s) = \eta(i) \\ G(x|s) = G(x|i) \end{cases} \tag{4}$$

4.2 Estimation

We estimate the model by maximum likelihood, obtaining contributions on durations $\{t_j\}_{j \in U,S,F,I}$ and labor incomes $\{x_j\}_{j \in S,F,I}$.

To derive the durations' contributions to the likelihood function, we define the hazard rates for each labor market state. The hazard rate out of unemployment and employment are, respectively:

$$h_u = \sum_{v=s,f,i} \lambda(v)\left[1 - G(x^*(v)|v)\right] \tag{5}$$

$$h_v = \eta(v) \tag{6}$$

Since no hazard rates depend on the duration in the state, all the durations follow a negative exponential distribution with parameter equal to the corresponding hazard rate.

To derive the labor incomes' contributions to the likelihood function, first consider that the labor income observed in the data only includes accepted job offers. Conditioning on the model's equilibrium, the accepted offers satisfy $x \geq x^*(v)$ for job type v, leading to the density of the observed labor income being $g_x(x|v) = \frac{g(x|v)}{1-G(x^*(v)|v)}$. In addition, to account for the possibility of labor income being measured with error, we assume that the observed labor income is $x^o = x \times \epsilon$, where $\epsilon \sim Q(\epsilon)$ is measurement error with $\mathbb{E}[\epsilon|x] = 1$. The density function of observed labor income is then given by:

$$g_x^o(x^o|v) = \int_{x^*(v)} \frac{1}{x} q\left(\frac{x^o}{x}\right) \frac{g(x|v)}{1 - G(x^*(v)|v)} dx \tag{7}$$

In conclusion, the logarithm of the likelihood function is:

$$\begin{aligned}
\mathcal{L}(\Theta) &= \sum_{j\in U} \left[\log\left(h_u \exp(-h_u t_j) \times u\right)\right] \qquad (8)\\
&+ \sum_{j\in S} \left[\log\left(\eta(s)\exp(-\eta(s)t_j \times \int_{x^*(s)} \frac{1}{x} q\left(\frac{x_j^o}{x}\right) \frac{g(x|s)}{1-G(x^*(s)|s)} dx \times s\right)\right]\\
&+ \sum_{j\in F} \left[\log\left(\eta(f)\exp(-\eta(f)t_j \times \int_{x^*(f)} \frac{1}{x} q\left(\frac{x_j^o}{x}\right) \frac{g(x|f)}{1-G(x^*(f)|f)} dx \times f\right)\right]\\
&+ \sum_{j\in I} \left[\log\left(\eta(i)\exp(-\eta(i)t_j \times \int_{x^*(i)} \frac{1}{x} q\left(\frac{x_j^o}{x}\right) \frac{g(x_j|i)}{1-G(x^*(i)|i)} dx \times i\right)\right]\\
\Theta &= \{\lambda(v), \eta(v), G(x|v)\}_{v=s,,f,i} \times \{b, \theta, Q(\epsilon)\}
\end{aligned}$$

where, consistently with the model's steady state equilibrium, each contributions has been weighted by the probability of being in the corresponding labor market state.

An important advantage of using maximum likelihood estimation in our application is that we can employ the Log-likelihood Ratio Test (LR) to directly test the null hypothesis expressed by the set of constraints (4).[6]

The model identification is relative straightforward. Durations provide direct information to identify the hazard rates. Labor incomes identify wage offers distributions as long as they belong to a recoverable parametric distribution (Flinn and Heckman, 1982). Following previous literature, we assume $G(x|v)$ are lognormal with parameters $\{\mu(v), \sigma(v)\}$. The flow utility of unemployment (b) and the discount rate (ρ) are jointly identified through the equilibrium equation (1). Therefore, we reparameterize the likelihood to estimate ρU directly and then set $\rho = 0.12$

[6]Recall that the Log-likelihood Ratio statistic is $LR = -2\left[\mathcal{L}(\Theta_0) - \mathcal{L}(\hat{\theta})\right]$, where $\mathcal{L}(\Theta_0)$ is the value of the log-likelihood function under the null hypothesis and $\mathcal{L}(\hat{\theta})$ is the value of the log-likelihood function of the unconstrained model.

– the recommended discount rate for Latin America (Moore et al., 2020) – to recover b. Finally, we set $\tau = 0.16$ to match the 2016 Colombian payroll contribution (Fernández and Villar, 2017) and recover θ from the condition $\theta = \tau \rho U$.

4.3 Results

Table 2 reports the estimated parameters. The column *Unrestricted* presents the unconstrained model: this is the model presented in Section 4 where all the parameters are allowed to be different across labor market states. The column *Restricted* presents a specification where we impose the set of constraints (4): this is a model where the parameters for the informal self-employed labor market state and the informal employee labor market state are constrained to be equal.

The Log-likelihood Ratio Test is presented at the bottom of the Table and clearly rejects the Restricted model against the Unrestricted one. Therefore, the null hypothesis of informal self-employment and informal employment as employee being behaviorally indistinguishable labor market states is strongly rejected. We have also estimated and tested two intermediate models, one where we impose equality only on the mobility parameters ($\lambda(s) = \lambda(i); \eta(s) = \eta(i)$) and one where we impose equality only in the offer distributions parameters ($\mu(s) = \mu(i); \sigma(s) = \sigma(i)$). Both models are clearly rejected against the Unrestricted model.[7]

Looking at the actual point estimates in conjunction with the implied values reported in Table 3, we observe a very large difference in the dispersion of the labor income distributions between the two informal states, with the self-employed's standard deviation being more than double the employee's one. This is the actual source of the differences coming from the wage offers distributions, not a difference in mean offers. In terms of mobility parameters, the source of the difference is in the termination rates, with the self-employed's termination rate being less than a third than the employee's one.

5 Conclusions

The paper performs both a parametric and non-parametric analysis to address a fundamental question in the growing literature using search models to study labor market informality: should informal self-employment and informal employment as employee be considered two different labor market states? Both the non-parametric and the parametric tests strongly reject the equality of the two labor market states, cautioning against aggregating them in a common "informality state"

[7]Complete results are available in Appendix Table A.2.

as done in important previous contributions in the literature.[8] The parametric model indicates that the source of the difference are the higher dispersion in informal self-employment income offers than in informal employee wage offers and the lower duration of informal employee jobs with respect to informal self-employment positions.

[8] Meghir et al. (2015) is an influential example. For additional references see footnote 1.

References

Albrecht, James, Lucas Navarro, and Susan Vroman, "The Effects of Labour Market Policies in an Economy with an Informal Sector," *Economic Journal*, 2009, *119* (2000), 1105–1129.

Almeida, Rita and Pedro Carneiro, "Enforcement of labor regulation and informality," *American Economic Journal: Applied Economics*, 2012, *4* (3), 64–89.

Bobba, Matteo, Luca Flabbi, and Santiago Levy, "Labor Market Search, Informality, and Schooling Investments," *International Economic Review*, 2022, *63* (1), 211–261.

_ , _ , _ , and Mauricio Tejada, "Labor Market Search, Informality, and on-the-Job Human Capital Accumulation," *Journal of Econometrics*, August 2021, *223* (2), 433–453.

Bosch, Mariano and Julen Esteban-Pretel, "Job Creation and Job Destruction in the Presence of Informal Markets," *Journal of Development Economics*, July 2012, *98* (2), 270–286.

Charlot, Olivier, Franck Malherbet, and Mustafa Ulus, "Efficiency in a search and matching economy with a competitive informal sector," *Economics Letters*, 2013, *118* (1), 192 – 194.

Dodge, Yadolah, ed., *Kolmogorov-Smirnov Test in The Concise Encyclopedia of Statistics*, Springer New York, 2008.

Esteban-Pretel, Julen and Sagiri Kitao, "Labor market policies in a dual economy," *Labour Economics*, 2021, *68*, 101956.

Fernández, Cristina and Leonardo Villar, "The Impact of Lowering the Payroll Tax on Informality in Colombia," *Economía*, 2017, *18* (1), 125–155.

Flinn, Christopher J and James J. Heckman, "New Methods for Analyzing Structural Models of Labor Force Dynamics," *Journal of Econometrics*, 1982, *18* (1), 115–168.

_ and James J Heckman, "Are Unemployment and out of the Labor Force Behaviorally Distinct Labor Force States?," *Journal of Labor Economics*, 1983, *1* (1), 28–42.

Granda, Catalina and Franz Hamann, "Informality , Saving and Wealth Inequality in Colombia," *IDB Working Papers*, 2015, *575*, 1–43.

Haanwinckel, Daniel and Rodrigo R Soares, "Workforce composition, productivity, and labour regulations in a compensating differentials theory of informality," *The Review of Economic Studies*, 2021, *88* (6), 2970–3010.

Kanbur, Ravi, "Conceptualising Informality: Regulation and Enforcement," *IZA Discussion Papers*, May 2009, *4186*, 1–18.

Kaplan, E. L. and Paul Meier, "Nonparametric Estimation from Incomplete Observations," *Journal of the American Statistical Association*, 1958, *53* (282), 457–481.

Kiefer, Nicholas M., "Economic Duration Data and Hazard Functions," *Journal of Economic Literature*, 1988, *26*, 646–679.

Meghir, Costas, Renata Narita, and Jean-Marc Robin, "Wages and Informality in Developing Countries," *American Economic Review*, April 2015, *105* (4), 1509–1546.

Moore, Mark A., Anthony E. Boardman, and Aidan R. Vining, "Social Discount Rates for Seventeen Latin American Countries: Theory and Parameter Estimation," *Public Finance Review*, January 2020, *48* (1), 43–71.

Narita, Renata, "Self-employment in developing countries: A search-equilibrium approach," *Review of Economic Dynamics*, 2020, *35*, 1–34.

Perry, Guillermo E., William F. Maloney, Omar S. Arias, Pablo Fajnzylber, Andrew D. Mason, and Jaime Saavedra-Chanduvi, *Informality : Exit and Exclusion* number 6730. In 'World Bank Publications.', The World Bank, 2007.

Ulyssea, Gabriel, "Firms, informality, and development: Theory and evidence from Brazil," *American Economic Review*, 2018, *108* (8), 2015–2047.

Table 1: Kolmogorov-Smirnov Test of Equality of Distributions

	Statistic	P-value
H_0 : Informal E duration distribution = Informal SE duration distribution:		
Values	0.3869	0.0000
H_0 : U duration distribution before Informal E = U duration distribution before Informal SE		
Values	0.0326	0.0177
H_0 : Informal E labor income distribution = Informal SE labor income distribution		
Values	0.1106	0.0000

NOTE: E denotes employees, SE self-employed, and U unemployed.

Table 2: Search Model Estimated Parameters

Parameters	Unrestricted			Restricted		
	Formal	Informal E	Informal SE	Formal	Informal E	Informal SE
	$v = f$	$v = i$	$v = s$	$v = f$	$v = i$	$v = s$
$\lambda(v)$	0.0897	0.0553	0.0529	0.097	0.0446	
	(0.0015)	(0.0012)	(0.0009)	(0.0015)	(0.0007)	
$\eta(v)$	0.0157	0.0317	0.0096	0.0159	0.0115	
	(0.0001)	(0.0006)	(0.0001)	(0.0001)	(0.0001)	
$\mu(v)$	0.2956	0.0071	-0.0500	0.2957	-0.0363	
	(0.0033)	(0.0055)	(0.0001)	(0.0068)	(0.0068)	
$\sigma(v)$	0.3145	0.3434	0.5048	0.3141	0.4723	
	(0.0099)	(0.0098)	(0.0058)	(0.0222)	(0.0142)	
b	-1.7205			-1.6053		
	(0.0839)			(0.0954)		
θ	0.0106			0.0160		
	(0.0048)			(0.0060)		
σ_ϵ	0.1196			0.1205		
	(0.0275)			(0.0509)		
Log-Likelihood	-470885.0			-480498.0		
LR Statistic	–			19226.0		
P-value	–			0.0000		

NOTE: Bootstrapped standard errors in parenthesis. E denotes employees, and SE self-employed. The Restricted Model imposes: $\lambda(s) = \lambda(i)$, $\eta(s) = \eta(i)$ and $\mu(s) = \mu(i)$, $\sigma(s) = \sigma(i)$. LR denotes the Log-likelihood Ratio Test.

Table 3: Search Model Implied Values

Values	Unrestricted			Restricted		
	Formal	Informal E	Informal SE	Formal	Informal E	Informal SE
	$v = f$	$v = i$	$v = s$	$v = f$	$v = i$	$v = s$
Employment:						
$E[t\|v]$	63.5	31.5	104.4	63.0	86.8	
$E[x\|v]$	1.412	1.068	1.080	1.412	1.078	
$SD[x\|v]$	0.207	0.143	0.339	0.207	0.291	
Unemployment:						
$E[t\|u]$		5.1			5.4	

NOTE: Values obtained from the point estimates reported in Table 2.

A Appendix

Table A.1: Descriptive Statistics

	Unemployed $v=u$	Formal $v=f$	Informal E $v=i$	Informal SE $v=s$
Duration (months):				
$E[t\|v]$	4.0	64.8	32.8	105.7
$SD[t\|v]$	6.9	76.6	53.1	103.9
Labor Income (US dollars per hour):				
$E[x\|v]$	–	1.42	1.06	1.07
$SD[x\|v]$	–	0.55	0.36	0.56
Labor Market States (Proportion of the Labor Force):				
v	0.09	0.40	0.12	0.39

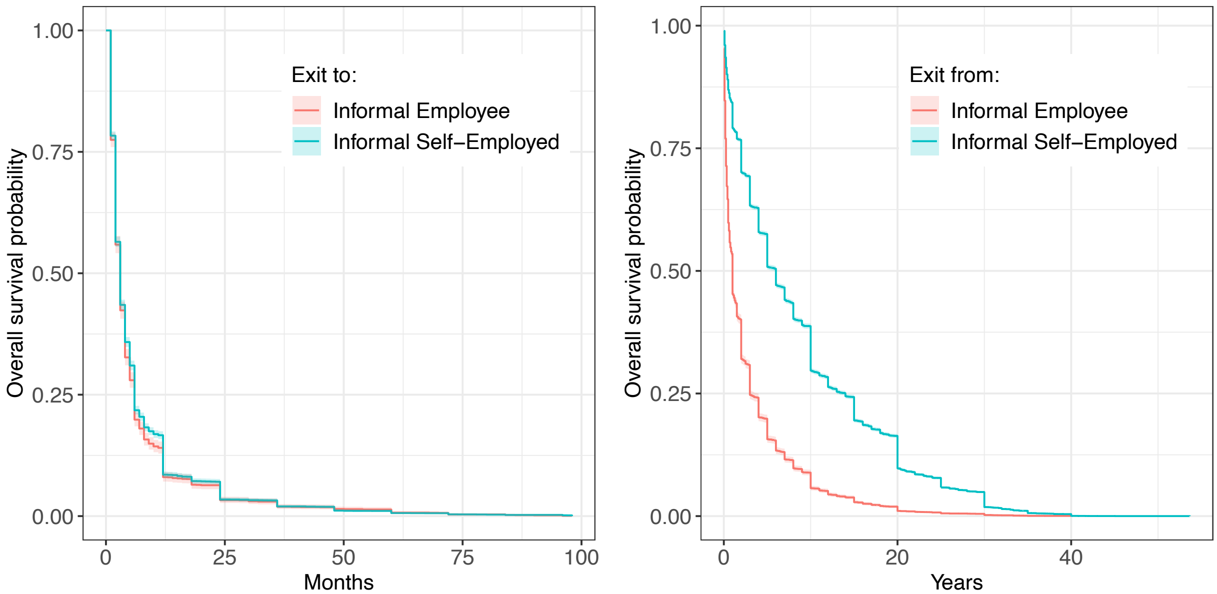
NOTE: The sample is comprised of 25-55 years old men, living in urban areas, that have completed at most secondary education and work full-time when employed

Table A.2: Estimated Parameters on the Unrestricted and Three Different Restricted Models

Parameter	Model 1	Model 2	Model 3	Model 4
$\lambda(s)$	0.0529	0.0446	0.0529	0.0446
	(0.0009)	(0.0065)	(0.0009)	(0.0007)
$\lambda(f)$	0.0897	0.097	0.0897	0.097
	(0.0015)	(0.0015)	(0.0015)	(0.0015)
$\lambda(i)$	0.0553	–	0.0553	–
	(0.0012)	–	(0.0012)	–
$\eta(s)$	0.0096	0.0115	0.0096	0.0115
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
$\eta(f)$	0.0157	0.0159	0.0157	0.0159
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
$\eta(i)$	0.0317	–	0.0317	–
	(0.0006)	–	(0.0006)	–
$\mu(s)$	-0.0500	-0.0500	-0.0418	-0.0363
	(0.0)	(0.0)	(0.0044)	(0.0068)
$\sigma(s)$	0.5048	0.5048	0.484	0.4723
	(0.0058)	(0.0333)	(0.009)	(0.0142)
$\mu(f)$	0.2956	0.2957	0.2902	0.2957
	(0.0033)	(0.0082)	(0.004)	(0.0068)
$\sigma(f)$	0.3145	0.3147	0.3314	0.3141
	(0.0099)	(0.0458)	(0.0126)	(0.0222)
$\mu(i)$	0.0071	0.007	–	–
	(0.0055)	(0.0806)	–	–
$\sigma(i)$	0.3434	0.3436	–	–
	(0.0098)	(0.0175)	–	–
x^*	0.0664	0.0734	0.0243	0.1000
	(0.0302)	(0.1081)	(0.0314)	(0.0374)
σ_ϵ	0.1196	0.1191	0.0578	0.1205
	(0.0275)	(0.0491)	(0.0386)	(0.0509)
Log-Likelihood	-470885.0	-479755.0	-471627.0	-480498.0
LR Statistic	–	17740.0	1484.0	19226.0
P-value	–	0.0000	0.0000	0.0000

NOTE: Bootstrapped standard errors in parenthesis. Model 1 is the unrestricted model. Model 2 impose the constraints $\lambda(s) = \lambda(i)$ and $\eta(s) = \eta(i)$. Model 3 impose the constraint $\mu(s) = \mu(i)$ and $\sigma(s) = \sigma(i)$. Finally, model 4 combine the constraints in model 2 and 3.

Figure A.1: Kaplan-Meier Survival Functions



(a) Unemployment Duration conditional on Exit

(b) Employment Duration

Figure A.2: Empirical CDF of accepted hourly wages